



## 저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

농학석사 학위논문

과거 병발생 자료를 이용한 배 검은별무늬병  
발생 예측모형 개발과 검증

**Development and evaluation of a pear scab  
forecast model using historical data**

2019년 8월

서울대학교 대학원

농생명공학부 식물미생물학전공

서 수 연

**A THESIS FOR THE DEGREE OF MASTER OF SCIENCE**

**Development and evaluation of a pear scab  
forecast model using historical data**

**BY**

**SOO YEON SEO**

**Department of Agricultural Biotechnology**

**The Graduate School of Seoul National University**

**August 2019**

## 초 록

식물병의 발생은 기상 조건에 따라 달라지며 날씨를 기반으로 하는 식물병 발생 예측은 농약 살포를 위한 의사결정에 유용하다. 식물병 방제 의사결정은 식물병 발생 위험 수준을 평가할 수 있는 위험요소의 임계값을 기준으로 결정된다. 본 연구에서는 6월 1일 이전에 배 검은별무늬병이 발생할 가능성을 예측하는 모형을 개발하고 검증하였다. 2014년 21개 지점과 2015년 18개 지점의 배 검은별무늬병 발생 자료와 기상자료를 분석하여 월동 후 6월 1일 이전에 병발생 가능성을 예측하는 모형을 개발하였다. 2016년부터 2018년까지 각각 76, 69, 57개의 지점의 병발생 및 기상자료를 분석하여 앞서 개발한 예측모형을 검증하였다. 예측모형 개발에는 사과 검은별무늬병 예측모형으로 잘 알려진 Mills 모형을 이용하여 계산된 일별 감염위험( $IR_d$ )을 위험요소 지표로 활용하였다. 또한 배 검은별무늬병의 역학적 특성과 과거 병발생 자료를 고려하여 월동 후 병원균의 감염기간(4월 1일 - 5월 16일)과 경제적 피해를 심각하게 줄 수 있는 초기 병발생 위험종료일(6월 1일)을 정하였다. 배 검은별무늬병 발생에 관한 감염위험의 임계값을 구하기 위하여 4월 1일부터 6월 16일까지  $IR_d > 0$ 인 날의 수를 적용하였다. Receiver operating characteristic (ROC) curve 분석을 이용하여 구한 감염 위험도의 최적 임계값은 18일이었다. 이 임계값으로 개발한 배 검은별무늬병 예측 모형을 2x2 contingency table 분석하였고 그 결과 경보적중률은 94 %, 오경보율은 17 %였으며 모형의 검증에서는 경보적중률과 오경보율이 각각 70 %와 73%이었다.

**주요어 :** 배 검은별무늬병, 식물병 방제, 예찰모형, 의사결정, 이항로지스틱회귀분석, ROC curve

**학 번 :** 2017-27751

# 목 차

	<i>page</i>
초 록 .....	i
목 차 .....	ii
표목차.....	iv
그림목차.....	v
I. 서 론 .....	1
II. 재료 및 방법.....	7
1. Mills 모델을 이용한 병발생 위험도 추정 .....	7
2. 배 검은별무늬 자료 및 초기 감염시기 추정 .....	10
3. 기상 관측 자료.....	13
4. 이항로지스틱회귀분석을 통한 병 발생확률 분석 .....	14
5. ROC curve 분석을 통한 감염위험수준 최적 임계값 추정 .....	15
6. 2016 년, 2017 년, 2018 년 자료를 이용한 모델 평가 .....	16

<b>III. 결 과</b> .....	<b>17</b>
1. 이병엽율을 통한 병발생 위험구간 추정.....	17
2. 이항로지스틱회귀분석 결과.....	19
3. ROC curve 분석 결과 .....	19
4. 감염위험 수준 최적 임계값 설정 .....	22
5. 최적 임계값 검증 .....	24
 <b>IV. 고 찰</b> .....	 <b>26</b>
 <b>V. 참고문헌</b> .....	 <b>30</b>
 <b>ABSTRACT</b> .....	 <b>34</b>

## 표 목 차

	<i>page</i>
<b>Table 1.</b> Two-way contingency table .....	6
<b>Table 2.</b> Two-way contingency table analysis for pear scab prediction in 2016- 2018 using the threshold of IRd=18 days .....	25

## 그림 목 차

	<i>page</i>
<b>Figure 1.</b> Locations of orchards where disease incidence of pear scab on leaves was monitored in Korea in 2014-2018 .....	12
<b>Figure 2.</b> Disease progress curves of pear scab at 39 locations in Korea in 2014 and 2015.....	18
<b>Figure 3.</b> Binary logistic regression curve showing probability of pear scab occurrence prior to June 1 versus number of days with infection risk of greater than zero.....	20
<b>Figure 4.</b> Receiver operating chracteristic (ROC) curve derived from the binary logistic regression analysis in Fig. 3.....	21
<b>Figure 5.</b> Probability of true positive warnings and false positive warnings on pear scab occurrence based on the number of days with infection risk of greater than 0 during the period of April 1 to May 16 in 2014 and 2015.....	23



## I. 서 론

농부들은 작물의 작기 전, 작기 중과 후에 농약 살포에 대한 끊임없는 의사 결정을 한다. 의사 결정에 있어 작물 손실 피해는 가장 크게 고려되는 사항이다. 그러나 의사 결정은 어떤 계획을 세우더라도 완벽하지 않으며 종종 맞지 않은 결과를 도출할 수 있다. 즉 농약을 살포했지만 실제로 필요하지 않거나, 농약 살포가 필요한데 살포하지 않아 병이 발생할 수 있다.

식물병을 방제하기 위한 농약은 기상 환경이나 작물의 상태와 상관없이 살포되어 농약이 남용되기도 한다(Gillespie and Sutton, 1979). 또한 많은 양의 농약은 심각한 환경 오염과 함께 인간의 건강에 해가 되며, 병해충의 농약 저항성을 높인다(Song et al., 2013; Fabre et al., 2003). 따라서 계획적인 농약 살포는 비용적, 환경적 측면에서 이득이 되며, 이때 통계학적인 접근은 효율적인 결과를 내는 데에 도움이 된다(Madden, 2006). 통계학적인 접근을 위한 식물 병 예측 모델들은 농약 살포 시기를 예측하는 것에 초점을 맞춘다. 이를 위해 병해충이 어떠한 손실을 일으키는지 오랜 시간 관찰하며 병해충을 일으키는 위험요소를 파악하여야 한다(Hughes, 1999).

식물병 발생 위험요소를 분석하는 연구는 여러 작물에서 이루어지고 있다. Gillespie and Sutton (1979)은 온도, 엽면수분, 강우 등 기상요소들의

분석하여 당근 잎마름병 병원균(*Alternaria dauci*)의 감염에 좋은 환경을 구명하였고, 이를 이용하여 효율적인 방제 시기를 추정하였다. Mila et al. (2005)는 온도, 강우, 관개 등이 피스타치오의 이삭마름병과 줄기마름병(*Fusicoccum* sp.)에 미치는 영향을 연구하여 병발생 위험을 평가하였다. Lindblad and Waern (2002)은 관행재배에 따른 밀 위축병(Wheat Dwarf Virus (WDV)) 발생 위험을 분석하여 병 방제법을 제시하였다. Neufeld et al. (2018)은 조롱박에서 노균병(*Pseudoperonospora cubensis*)을 일으키는 병원균의 감염 위험 모델을 평가하고 병 방제 비용과 농작물 피해를 최소화하는 임계값을 추출하였다. 이처럼 병발생 가능성을 판단할 수 있는 위험요소에 대한 연구는 미래의 병발생을 예측하여 효율적인 방제 의사결정을 하는 데에 사용된다.

배는 우리나라의 전체 과실 재배 면적 중 6 %(103.03 km<sup>2</sup>)를 차지하며 연간 20만톤 이상이 생산된다(농산물생산조사, 통계청, 2018). *Venturia nashicola*에 의해 발생하는 배 검은별무늬병은 배나무에서 발생하는 병해충 중 피해가 극심하고 관리가 어려운 식물병이며, 우리나라 배 재배에 있어 경제적으로 가장 큰 피해를 주는 병해 중 하나이다(Song and Seo, 2018; Nam et al., 2014). 배 검은별무늬병은 이른 봄부터 늦은 가을까지 작기 전체에 나타나며 과실과 잎, 어린 줄기를 감염시키고 배 생산성을 크게 줄인다(Li et al., 2005). 이러한 배 검은별무늬병을 방제하기 위해 농약

살포가 필요하기에 연간 농약 살포 횟수는 평균 10회 수준의 많은 농약으로 이루어진다(Song et al., 2013). 따라서, 최소한의 농약 살포로 배 검은별무늬병 발생을 효율적으로 방제하기 위해서는 병에 대한 정확한 위험요소 분석이 필요하다.

Mills 모델은 사과 검은별무늬병(*Venturia inaequalis*) 감염 위험요소를 분석하여 식물병 감염 위험수준을 계산한다. 이 모델은 병발생과 온도, 엽면수분의 상관관계를 기반으로 온도에 따라 병발생에 필요한 최소한의 엽면수분지속시간을 제시한다(MacHardy and Gadoury, 1989). 검은별무늬병 발생생태는 사과와 배가 유사하기 때문에 Mills 모델은 배 검은별무늬병 발생예측에도 사용되기도 한다(Li et al., 2007).

농약 살포 비용을 최소화하기 위한 의사결정에는 농약 살포의 기준이 되는 위험수준의 임계값이 필요하다(Neufeld et al., 2018). 이항로지스틱 회귀분석(Binary logistic regression analysis)은 경험적 자료를 바탕으로 감염 위험요소의 중요성을 평가하고, 위험요소의 수준에 따라 식물병 발생 확률을 추정하는 분석방법으로 사용된다(Yuen et al., 1996).

이항로지스틱회귀분석으로 구한 병 발생확률을 평가하는 방법으로 Receiver Operating Characteristic (ROC) curve 분석이 있다(Pethybridge, 2009). ROC curve 분석은 예측값을 평가하거나 작물 관리 비용을 최소화할 수 있는 적절한 병발생 위험수준의 임계값을 찾을 때 쓰이는 통계학적인 방법이다(Pethybridge et al., 2009). 병발생 위험수준의

적절한 임계값을 찾기 위해서는 위험수준을 임의의 기준으로 나누어서 판단한다(Table 1). 위험수준으로 예측한 자료는 Predicted condition으로 실제 병 발생을 관찰한 자료는 Observed condition으로 분류한다. Predicted condition에서 임의의 병 감염 위험수준을 기준으로 이 값을 초과하는 자료는 병이 발생( $P^+$ )할 것이라 예측하고 기준값과 동일하거나 낮은 자료는 병이 발생하지 않을( $P^-$ ) 것이라 예측한다. 분류한 두 예측값은 실제 병 발생을 관찰한 자료를 기준으로 병 발생( $O^+$ )과 병 미발생( $O^-$ )으로 나뉜다. 이때 기준이 되는 위험수준이 낮게 설정되면 실제로 병이 발생할 상황을 예측하는 능력은 향상되지만 실제로 병이 발생하지 않은 경우 또한 병이 발생한 것으로 예측한다. 이때의 오류를 Type 1 오류 또는 False Positive라고 한다. 이를 오경보율이라고도 하며 병 미발생 경우( $O^-$ ) 중 병발생 예측값( $P^+$ )의 비로 나타낸다. 반면에 위험수준이 높게 설정되면 병이 발생하지 않은 경우를 예측하는 능력은 향상되지만 병이 발생한 경우 또한 병이 발생하지 않은 것으로 예측하기도 한다. 이때의 오류는 Type 2 오류 또는 False Negative라고 하며 실제 병발생 경우( $O^+$ ) 중 병이 발생하지 않을 것이라 예측한 경우( $P^-$ )의 비로 나타낸다(Madden, 2006). 따라서 두 가지 오류를 최소화하기 위한 적절한 기준의 설정은 중요하며 이때의 기준을 최적 임계값이라 한다(Lee et al., 2017). 최적 임계값을 기준으로 True positive와 True negative를 각각 실제 병 발생한 상황( $O^+$ ) 중 병이 발생할 것이라 예측한 경우( $P^+$ )와 실제 병이 발생하지 않은 상황( $O^-$ )에서 병이

발생하지 않을 것이라 예측한 경우( $P^-$ )로 정확성을 판단한다.

본 연구에서는 배 검은별무늬병의 이병엽을 자료를 수집하고, 병발생 확률을 이항로지스틱회귀분석과 ROC curve 분석을 병발생 위험요소의 임계값을 설정하고, 이 임계값을 이용하여 병발생 위험 예측의 정확성을 평가하고자 한다.

Table 1. Two-way contingency table

		Observed condition	
		Positive ( $O^+$ )	Negative ( $O^-$ )
Predicted condition	Positive ( $P^+$ )	True positive (Sensitivity <sup>a</sup> )	False positive (Type 1 error)
	Negative ( $P^-$ )	False negative (Type 2 error)	True negative (Specificity <sup>b</sup> )

<sup>a</sup> Sensitivity =  $(O^+ \cap P^+)/O^+$

<sup>b</sup> Specificity =  $(O^- \cap P^-)/O^-$

## II. 재료 및 방법

### 1. Mills 모델을 이용한 병발생 위험도 추정

이 연구에서는 사과 검은별무늬병 감염 예측모델인 Mills 모델(Eq. 1)을 활용하였다. Mills 모델은 특정 온도( $T_t$ ) 조건에서 사과 검은별무늬병 감염에 필요한 최소한의 엽면수분지속시간( $MW_t$ )을 다음과 같이 정의한다(MacHardy and Gadoury, 1989).

$$MW_t = 86.3749 - 18.0034T_t + 1.5703T_t^2 - 0.0062T_t^3 + 0.0009198T_t^4$$

(Eq. 1)

여기서 입력값으로 자동기상관측기(Automatic Weather Station; AWS)로부터 생성되는 매시간( $t$ ) 온도 자료를 사용하므로 최소엽면수분지속시간( $MW_t$ )은 시간별 자료로 구해진다.

배 검은별무늬병의 시간별 감염위험도( $IR_t$ )는 Mills 모델로 계산한 최소엽면수분시간( $MW_t$ ) 대비 관측된 엽면수분지속시간( $W_t$ )으로 계산하였다. 매시간 엽면수분지속시간( $W_t$ )은 해당 시간 중 잎이 젖어 있는 시간의 비율로 나타내므로  $0 \leq W_t \leq 1$ 의 범위를 갖는다. 즉, 한 시간 동안 잎이 말라 있으면  $W_t = 0$ 이며, 젖어 있으면  $W_t = 1$ , 30분 동안만 젖어 있으면  $W_t = 0.5$ 이다. 따라서 시간별 감염위험도는 다음과 같이 계산된다(Eq. 2).

$$IR_t = \frac{W_t}{MW_t}$$

(Eq. 2)

일별 감염위험도(  $IR_d$  )는 시간별 감염위험도(  $IR_t$  )의 누적값( $ACCIR_t$ )으로 계산한다. 기주 식물의 표면이 장시간 건조해질 경우, 이미 발아된 병원균 포자는 생존이 불가능하게 된다. 예컨대 사과 겹무늬썩음병균(*Botryosphaeria dothidea*)의 경우 12시간 이상 건조 상태가 지속되면 병원균 포자의 활성이 회복되지 못한다(Kim et al., 2016). 따라서, 건조 시간이 12시간 이상 지속될 경우, 그 시간까지 누적된 감염위험도( $ACCIR_t$ )를 '0' 에서 재시작 하도록 하였다. 또한 일정량 이상 비가 오게 되면 빗물에 의해 기주 식물 표면의 포자가 씻겨 나가 감염률이 떨어진다고(Suzuki, 1975). 이 연구에서는 시간별 강수량이 4 mm 이상일 경우에도 누적감염위험도( $ACCIR_t$ )를 '0' 에서 재시작 하도록 설정하였다. 즉, 건조 지속기간이 12시간보다 길지 않고, 시간별 강수량이 4 mm 보다 적을 때, 시간별 감염위험도(  $IR_t$  )는 매시간 누적되어 누적감염위험도( $ACCIR_t$ )를 산출한다 (Eq. 3).

$$ACCIR_t = \begin{cases} ACCIR_{t-1} + IR_t & \text{if rain} < 4 \text{ mm/hr} \\ & \text{and accumulated dry hour} \leq 12 \text{ hr at time } t \\ 0 & \text{if otherwise} \end{cases}$$

(Eq. 3)



따라서 특정 날짜의 감염위험도(  $IR_d$  )는 당일 매시간 누적된 감염위험도( $ACCIR_t$ )의 최대값으로 다음(Eq. 4)과 같이 결정하였다.

$$IR_d = Max(ACCIR_t)$$

(Eq. 4)

## 2. 배 검은별무늬병 발생 자료 및 초기 감염시기 추정

배 검은별무늬병 병원균의 감염시기를 추정하기 위하여 우리나라의 배 검은별무늬병 이병엽율 자료를 농촌진흥청 국가농작물병해충관리시스템(National Crop Pest Management System; NCPMS)으로부터 수집하였다. 이 자료는 각 지역별 농업기술센터의 관찰포에서 조사된 자료로 2014년부터 2018년까지 연도별로 각각 21, 18, 76, 69 그리고 57개의 지점에 위치한 배과수원에서 조사된 자료이다. 이 중 2014년과 2015년의 자료를 일일 감염위험도 임계값 설정에 사용하였고, 2016년, 2017년, 2018년 자료를 감염위험도 임계값을 이용한 병발생 위험 예측모형 검증에 사용하였다(Figure 1).

병발생 조사는 4월부터 10월까지 매월 1일과 16일에 이루어졌다. 배 검은별무늬병의 조사 방법은 농촌진흥청에서 제공하는 농작물병해충예찰요강(농촌진흥청, 2016)에 따라 대상 과원에서 5그루를 선정하여 매회 동일 나무당 500잎의 병든잎수를 조사하였다. 그 결과를 토대로 5그루의 평균 이병엽율을 다음과 같은 식(Eq. 5)을 이용하여 구하였다.

$$\text{이병엽율}(\%) = \frac{\text{병든잎수}}{\text{조사잎수}} \times 100$$

(Eq. 5)

수집한 이병엽을 자료를 바탕으로 배 검은별무늬병의 초기감염위험시기를 추정하였다. 각 지점에서 배 생육기간 동안 이병엽율의 변화를 확인한 결과 전국적으로 매년 월동 후 검은별무늬병이 대체로 5월 1일부터 6월 1일 경에 처음 관찰되었음을 알 수 있었다(Figure 2). 또한 6월 1일 이후에도 처음 병징이 관찰된 경우도 있었으나 이 경우에는 병 진전이 크게 진행되지 않아 이로 인한 경제적 피해는 미미할 것으로 판단되었다. 따라서 이 연구에서는 월동 후 4월 말부터 6월 1일까지 검은별무늬병이 발생할 경우 배 생산성에 경제적인 피해가 발생할 것으로 가정하였다. 월동 후 병원균의 초기감염위험시기는 병 조사가 시작된 4월 1일부터 시작하여 경제적 피해를 줄 수 있는 최종 병발생일인 6월 1일을 기준으로 병원균의 잠복기간을 고려하여 추정하였다. 배 검은별무늬병의 병원균인 *Venturia nashicola*의 잠복기는 2~3주로 보고된 바가 있다(Umemoto, 1991, Li et al., 2007). 따라서 초기감염위험시기를 4월 1일부터 초기 병발생 위험 종료일(6월 1일)의 2주 전(5월 16일)까지의 기간(총 46일)으로 정하였다.

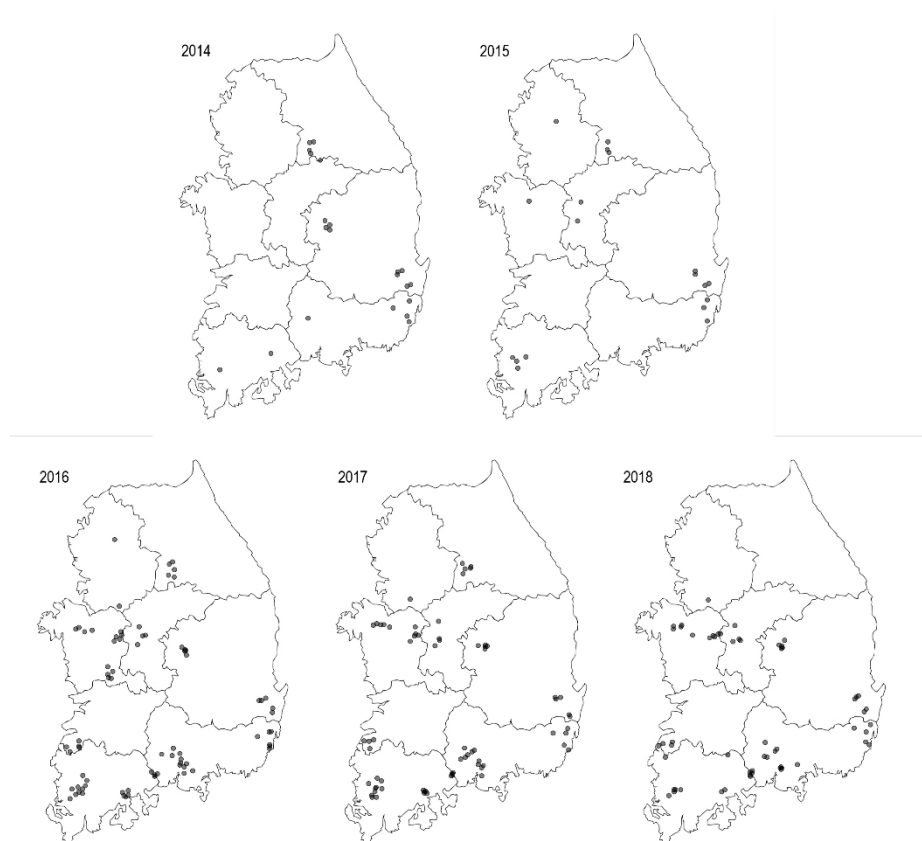


Figure 1. Location of orchards where disease incidence of pear scab on leaves was monitored in Korea in 2014-2018.

### 3. 기상 관측 자료

초기감염위험시기의 일별위험수준을 구하기 위하여 기상관측 자료를 기상청으로부터 수집하였다. 2014년부터 2018년까지 배 검은별무늬병 이병엽율이 조사된 241개 지점으로부터 최인접한 지점에 설치된 기상청 자동기상관측기(AWS)에서 측정한 관측자료를 사용하였다. 병발생 조사 지점과 최인접 기상관측 지점 간의 거리는 0.67 ~ 9.69 km였다. 관측되는 기상 자료 중 온도(℃), 습윤시간(hr), 강우량(mm)과 상대습도(%)를 사용하였다. 기상자료는 60초 단위로 측정되며 매시간 자료로 가공하여 출력된다. 습윤시간이 측정되지 않은 지점의 경우, 상대습도가 85 % 이상이거나 강우량이 0.1 mm 이상일 때, 식물체가 젖어 있는 것으로 판단하였다.

#### 4. 이항로지스틱회귀분석을 통한 병 발생확률 분석

배 검은별무늬병 발생 위험수준에 따른 병 발생확률을 구하기 위해 이항로지스틱회귀분석(Binary Logistic Regression Analysis)을 R programming language를 이용하여 수행하였다(R Core Team, 2017, Harrell Jr, 2019). 이 때 이항로지스틱 모델의 독립변수로 Mills 모델을 이용해서 계산한 지점별 검은별무늬병의 월동 후 초기 잠복기간(4월 1일 ~ 5월 16일) 동안의 일별 감염위험수준이 0 이상이었던 날의 수(IRd)를 활용하였다. 즉 6월 1일 이전에 병원균의 감염이 가능한 날의 수를 독립변수로 사용하였다. 종속변수로 6월 1일까지 병발생 유무에 따라 병이 발생한 지점은 '1', 발생하지 않은 지점은 '0'으로 값을 배정하였다. 감염 가능한 날의 수(IRd)에 따른 배 검은별무늬병 발생 확률( $p$ )은 다음(Eq. 6)의 이항로지스틱 모델을 이용하여 구하였다.

$$p = \frac{1}{1 + e^{-f(\text{IRd})}} \quad (6)$$

여기서  $f(\text{IRd}) = b_0 + b_1 \times \text{IRd}$  이다. 모델 파라미터인  $b_0$  와  $b_1$  을 추정하기 위하여 R 프로그램으로 모델 fitting을 수행하였다.

## 5. ROC curve 분석을 통한 최적 임계값 추정

이항로지스틱회귀분석을 통하여 구한 배 검은별무늬병 발생확률을 이용하여 ROC curve 분석을 수행하였다. ROC curve 분석은 R programming language를 통해 이루어졌으며 Epi, pROC, ROCR, Verification package를 사용하였다(R Core Team, 2017; Carstensen et al., 2018; Robin et al., 2011; Sing et al., 2005; Mason and Graham, 2002). 이 연구에서 ROC curve는 병발생 확률의 변화에 따른 민감도와 특이도를 그래프로 나타낸 것으로 curve의 아래 면적(AUC; Area Under Curve)은 병발생 예측의 정확성을 나타낸다. ROC curve은 (1, 0)에 가까울수록 정확도가 높으며, (1, 0)에 가까울수록 정확도가 낮다. 일반적으로 AUC값이 1에 가까워질수록 분석의 정확도가 높음을 뜻하고, 0.6 이하일 때는 분석의 의미가 없다고 본다(Muller et al., 2005; Pethybridge et al., 2009). AUC가 0.5일 경우에는 무작위(random)로 병발생을 예측하는 것과 마찬가지이다.

민감도는 병이 발생할 것이라 예측을 하고, 실제 병이 발생한 비율로 정보적중률을 나타낸다. 오경보율은 병이 발생할 것이라 예측하였으나 실제로 발생하지 않았을 때를 뜻하며 ‘1-특이도’ 값으로 구한다. 정보적중률과 오경보율의 차이가 최대값일 때의 감염위험도를 최적 임계값으로 결정하였다(Youden, 1950).

## 6. 2016년, 2017년, 2018년 자료를 이용한 모델 평가

2014년과 2015년의 배 검은별무늬병 이병엽을 자료와 Mills 모델을 사용하여 구한 감염위험도의 최적 임계값을 2016년부터 2018년까지의 자료로 평가하였다. 2014년과 2015년과 동일한 방법으로 이병엽과 기상 관측 자료를 수집하였으며 2016년, 2017년과 2018년의 자료를 사용하였다. 모델의 평가는 2x2 contingency table을 통해 이루어졌다(Table 1). 총 202개의 지점을 병발생 위험 종료일인 6월 1일 이전에 병이 발생한 지점은 병발생( $O^+$ )으로 나누었다. 6월 1일 이후 병이 발생했거나 혹은 조사 기간동안 병이 발생하지 않은 지점은 병 미발생( $O^-$ )으로 분류되었다. 각각의 경우에서 구한 최적 임계값을 기준으로 임계값을 초과하면 병이 발생할 것이라 예측( $P^+$ )한 것으로 임계값과 동일하거나 그 값에 미치지 못하면 병이 발생하지 않을 것이라 예측( $P^-$ )한 것으로 분류하였으며 이를 통하여 정확도와 경보적중률, 오경보율을 분석하였다.



### Ⅲ. 결 과

#### 1. 이병엽율을 통한 병발생 위험구간 추정

배 검은별무늬병 이병엽율 자료 분석 결과 2014년 21개 지점과 2015년 18개 지점 중 각각 12, 16개 지점에서 6월 1일 이전에 병이 발생하였다. 배 검은별무늬병이 처음 관찰된 날짜는 5월 1일이고, 6개 지점에서 이날 처음 병이 발생하였다. 이병엽율의 최대값은 2015년 10 %로 경상북도 경주시에서 관찰되었다. 전체적으로 2014년 대비 2015년의 이병엽율이 높으며 병발생 지점 수 또한 많은 것으로 나타났다.

2014년과 2015년의 배 검은별무늬병의 병 진전 곡선을 6월 1일을 기준으로 나타냈다(Figure 2). 6월 1일을 기준으로 6월 1일 이전에 병이 발생했을 경우 빨간색으로, 병이 발생하지 않거나 그 이후에 병이 발생했을 경우에는 파란색으로 나타냈다. 그 결과 6월 1일 이전에 병이 발생한 지점의 경우, 비교적 병 진전이 최대 10 %까지 발생하였으며 최대 이병엽율이 4 %가 넘는 곳도 4지점 확인할 수 있었다. 그러나 6월 1일 이후에 병이 발생한 지점의 경우, 최대 이병엽율이 2 %에 미치지 못하는 것으로 나타났다. 따라서, 6월 1일 이전에 배 검은별무늬병이 발병하면 그 피해가 상대적으로 큰 것을 확인할 수 있었다.

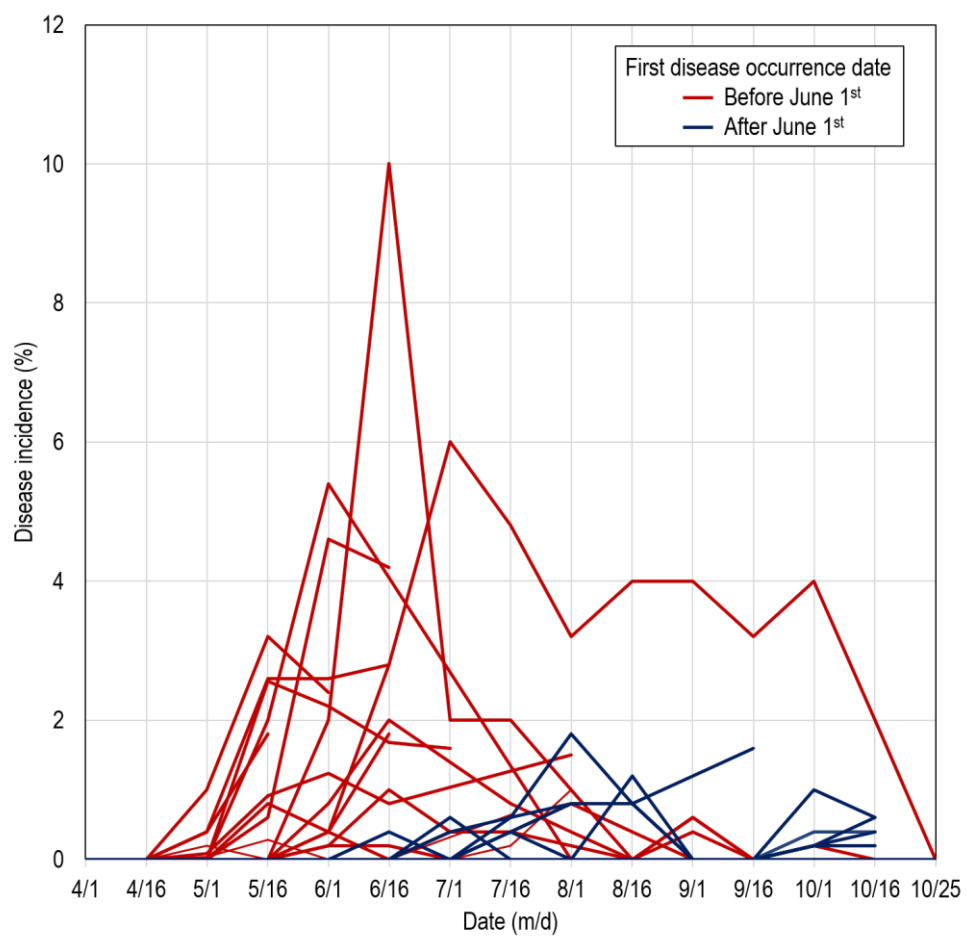


Figure 2. Disease progress curves of pear scab at 39 locations in Korea in 2014 and 2015

## 2. 이항로지스틱회귀분석 결과

2014년과 2015년의 배 검은별무늬병 발생 자료와 감염 위험 일 수 자료로 회귀계수  $b_0$  와  $b_1$  을 추정하였으며, 이 때, 구축된 이항로지스틱회귀 모형은 다음 Eq. 8과 같다.

$$p = \frac{1}{1 + e^{-( -10.82 + 0.54x )}}$$

(Eq. 8)

여기서  $x$  는 앞서 구한 IRd, 즉 병원군의 감염위험이 높을 것으로 판단한 날의 수를 나타낸다(Figure 3). 분석 결과 배 검은별무늬병 발생 확률은 IRd ( $r^2=0.667$ ,  $P<0.001$ )과 유의한 상관관계를 보였다. 만약 IRd가 16일이면 병 발생확률( $p$ )은 0.10, 18일이면 0.24, 20일이면 0.50이 된다.

## 3. ROC curve 분석 결과

이항로지스틱회귀분석 결과를 이용하여 적합성을 평가하기 위하여 ROC curve 분석을 진행하였으며 배 검은별무늬병 발생 예측에 대한 민감도와 특이도의 관계를 ROC curve로 나타냈다(Figure 4). 이 그래프에서 AUC 값은 0.93(95% CI, 85.45 %–100%;  $p<0.001$ )으로 분석의 정확도가 높은 것으로 나타났다.

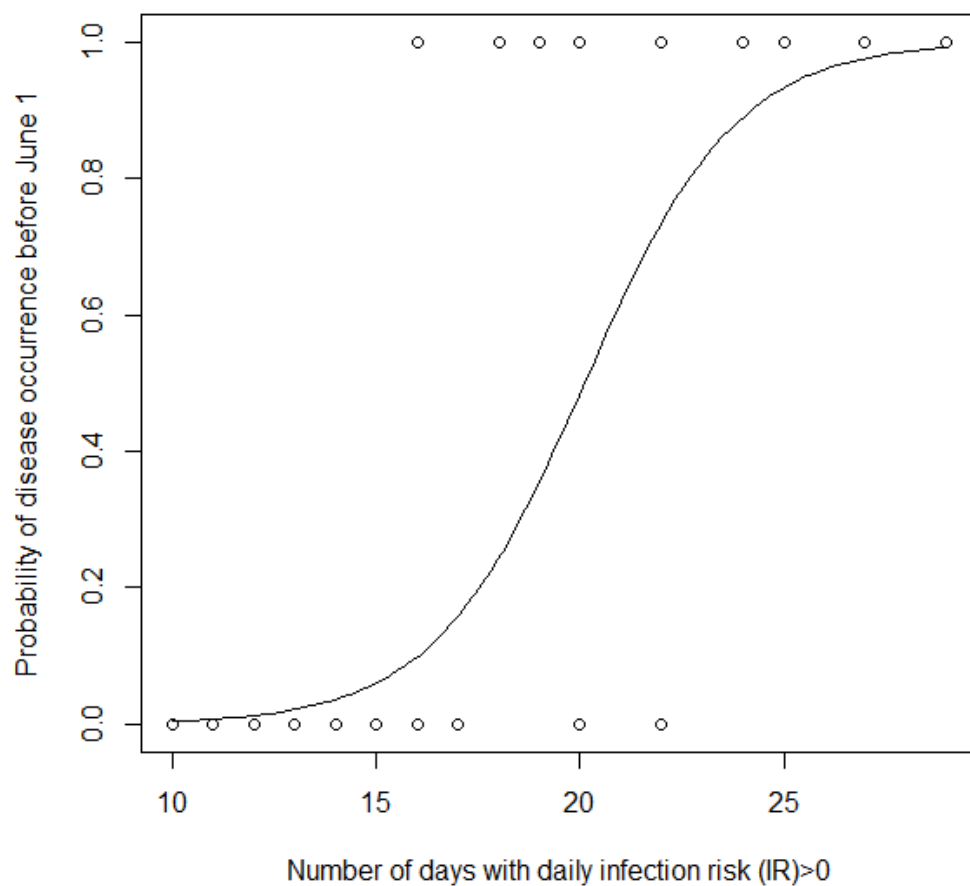


Figure 3. Binary logistic regression curve showing probability of pear scab occurrence prior to June 1 versus number of days with infection risk of greater than zero. The binary logistic regression analysis was conducted using the pear scab occurrence and weather data from 39 locations in Korea in 2014 and 2015.

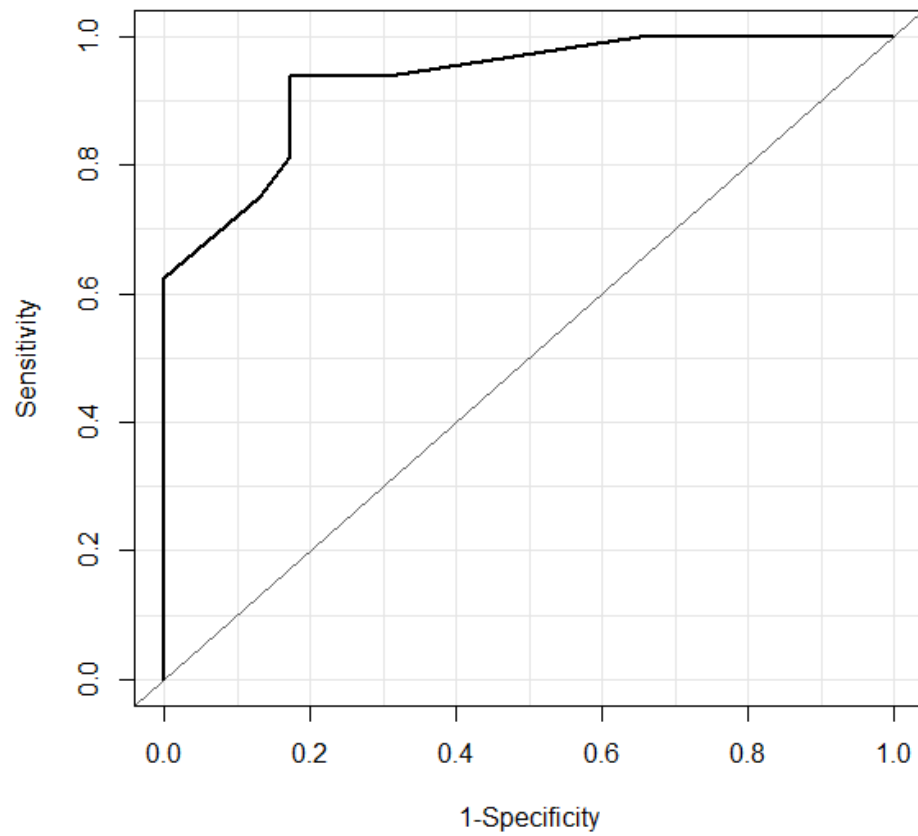


Figure 4. Receiver operating characteristic (ROC) curve derived from the binary logistic regression analysis in Fig. 3.

#### 4. 감염위험 수준의 최적 임계값 설정

감염위험 수준의 최적 임계값을 찾기 위하여  
경보적중률(sensitivity)과 오경보율( $1 - \text{specificity}$ ) 그래프를 Figure 5에  
나타냈다. 경보적중률과 오경보율의 차이가 최대값일 때 감염위험이 0보다  
큰 날 수(IRd)가 18일로 나타났다. IRd=18일을 감염위험 임계값으로 할 때,  
병이 발생한 16개의 지점 중, 경보 적중 지점은 15개 지점이고, 병이  
발생하지 않은 23개의 지점 중, 경보가 발생한은 지점은 4개 지점이었다.  
따라서 이 때의 경보적중률은 94 %이고 오경보율은 17 %로 나타났다.

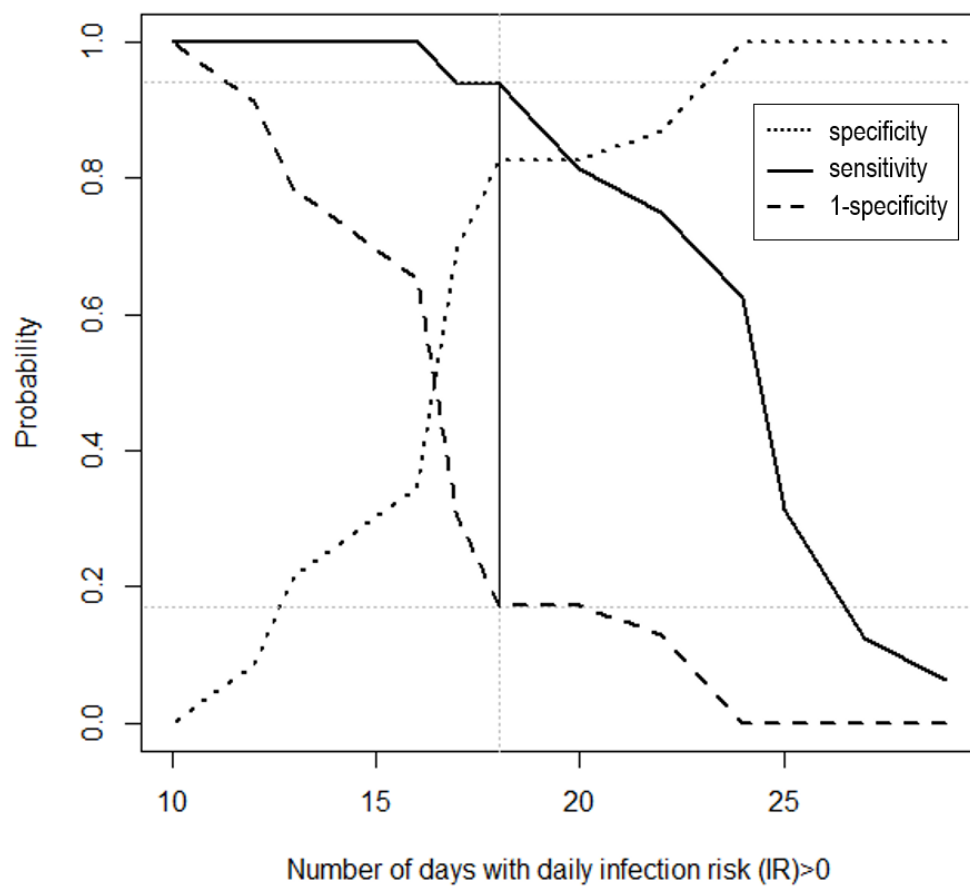


Figure 5. Probability of true positive warnings and false positive warnings on pear scab occurrence based on the number of days with infection risk of greater than 0 during the period of April 1 to May 16 in 2014 and 2015.

## 5. 최적 임계값 검증

2016년도부터 2018년까지 총 202개의 지점 중, 6월 1일까지 병이 발생한 지점은 84곳이었고, 병이 발생하지 않은 지점은 118지점이었다. 각 연도별로 4월 1일부터 5월 16일까지 기상값으로 추출한 일별 감염위험수준( $IR_d$ )이 0 이상인 날의 수(IRd), 즉 경보가 발생한 날의 수를 세었다. 경보 발생일을 18일을 기준으로 나누어 18일 이상 경보가 발생하면 6월 1일까지 병이 발생하는 것이라 예측하고, 그 미만이면 병이 발생하지 않을 것이라 예측하였다. 그 결과, 병이 발생할 것이라 예측한 지점은 145 지점, 병이 발생하지 않을 것이라 예측한 지점은 57 지점이었다. 이 결과를 바탕으로 two-way contingency table을 작성하였다(Table 2). 그 결과, 전체 지점 수 대비 sensitivity와 specificity를 합한 비율로 나타내는 정확도(accuracy)는 45 %이었다. 경보적중률은 70 %였으며 오경보율은 73 %이었다.



Table 2. Two-way contingency table analysis for pear scab prediction in 2016-2018 using the threshold of IRd=18 days

		Observed condition	
		Positive (O <sup>+</sup> )	Negative (O <sup>-</sup> )
Predicted condition	Positive (P <sup>+</sup> )	59	86
	Negative (P <sup>-</sup> )	25	32

## IV. 고 찰

본 연구는 Mills 모델을 사용하여 배 검은별무늬병의 일별 감염위험수준을 구하고 수집한 이병엽을 자료로 로지스틱 회귀분석을 수행하여 병 발생확률을 구하였고 회귀분석을 통해 구한 병 발생확률을 ROC curve 분석을 통해 평가하였다. 2014년과 2015년의 배 검은별무늬병 이병엽을 자료와 일별 감염위험수준의 분석 결과 4월 1일부터 경보가 발생한 일 수 즉, 일별 감염위험수준이 ‘0’을 넘는 값을 가지는 날이 18일 이상일 때, 병이 6월 1일 이전에 발생할 것이라 보았다. 이 때, 0.9가 넘는 AUC값을 가지므로 분석의 정확성 또한 높은 것으로 나타났다. 그러나 이 모형을 2016년부터 2018년까지의 자료로 검증하자 결과가 상이하게 나타났다. Two-way contingency table 분석으로 87 %의 정확도와 94 %의 경보 적중률, 17 %의 오경보율을 보였던 2014년과 2015년 분석 자료에 대비해 2016년부터 2018년의 검증 자료는 45 %의 정확도 밖에 보이지 않았으며 경보적중률은 70 %로 다소 높았으나 오경보율 또한 73%로 높았다. 이는 병이 발생할 것이라 예측한 지점에서 실제로는 병이 6월 1일 이전에 발생하지 않았음을 뜻한다. 즉 분석 결과에 따라 농약을 살포할 경우 불필요한 농약이 사용될 수 있음을 나타낸다. 이처럼 오경보율이 높게 측정된 이유는 다음과 같이 추정할 수 있다.

먼저 조사 방법에 따른 오류가 있을 수 있다. 식물병 발생 예찰에 있어 조사 시기는 매우 중요하다(Turechek and Wilcox, 2005). 배 검은별무늬병의 잠복기는 2주에서 3주로 알려져 있기 때문에 매월 1일과 16일에 조사된 자료의 경우, 병이 발생한 시기에 대한 정확한 자료를 얻을 수 없다. 예를 들어 2일에 병이 발생한 경우, 1일 조사에서 확인되지 않고 다음 조사일에 확인할 수 있으며 이때의 병 발생은 16일에 확인된다. 또한 조사일 후에 병이 발생하였으나 조기 낙엽으로 인해 병반 확인이 되지 않을 경우 다음 조사일에 병이 발생하지 않음으로 기록된다. 따라서, 매월 1일과 16일에 수집한 배 검은별무늬병 이병엽을 자료에서 조사 간격인 2일부터 15일 사이에 병이 발생했으나 조기 낙엽으로 인하여 이병엽이 제거되어 다음 조사일에 병이 발생하지 않은 것으로 조사되었을 수 있다. 즉, 실제로는 병이 발생했으나 병이 발생하지 않은 것으로 조사되었을 가능성이 있다.

배 검은별무늬병 이병엽율이 조사된 포장은 방제를 하지 않는 포장인 아닌 농약을 사용하는 포장으로서 약제 살포로 인한 병발생 억제가 있었을 것이다. 따라서, 병이 발생할 환경이 주어졌으나 농약 살포로 인하여 병이 발생하지 않았을 것이다.

배 검은별무늬병의 발생 확률 분석은 기상 자료의 분석으로 이루어졌다. 이는 병 조사가 이루어진 포장으로부터 가장 가까이 있는 AWS 관측기로부터 기상 자료를 수집하여 이루어졌다. 분석의 정확성을 위하여

포장으로부터 AWS까지의 거리가 10 km 이상인 지점은 제외하였으나 2014년부터 2018년까지 각 연도별로 평균 AWS 거리는 5.18, 5.20, 6.03, 6.15, 6.12 km로 평균 5.73 km이었다. 따라서, 가장 가까운 AWS 자료를 활용하였으나 그 거리에 따른 오차가 발생하여 실제 포장에서의 기상값과 차이가 있었을 가능성도 있다고 생각된다.

농약 살포에 대한 의사결정 도구 개발을 위한 위험수준의 파악에는 작물 손실과 병원균 감염 사이의 전제조건 분석이 필요하다(Fabre et al., 2003). 이 때의 위험수준의 파악은 방대한 경험과 주관적인 판단을 기본으로 하기 때문에 불분명한 요소가 존재하며 동일하지 않은 기후, 기주 식물 혹은 품종 같은 요인에 따라 달라질 수 있다(Yuen et al., 1996). 조사된 이병엽율의 자료는 배 과수원에서 동일하게 이루어졌으나 그 품종까지 분석에 고려하지 않았다. 따라서, 기주별로 감염 위험 요인이 미치는 영향이 다를 수 있기에 동일한 분석 적용시 정확도가 떨어질 수 있다(Neufeld et al., 2018).

본 연구에서는 sensitivity와 specificity를 동일한 가치를 두고 감염위험 수준의 임계점을 결정하였다. 그러나 농부들은 sensitivity와 specificity 두 값을 동일하게 취급하지 않는다(Pethybridge et al., 2009). 경제적 조건이나 자금 혹은 개인적인 목표와 같은 개인이 중요하게 여기는 요소들에 따라 두 값에 매겨지는 가치는 달라진다(Turechek and Wilcox, 2005). 게다가 농부들은 위험을 회피하려는 경향이 있어, 병의 발생을 놓치는 것보다 농약을 과다 살포하는 것을 선호한다(Neufeld et al., 2018;

Turechek and Wilcox, 2005). 따라서 본 연구는 병원균 감염위험수준을 기준으로 배 검은별무늬병에 의한 피해 발생 가능성 여부에 대한 정보를 제공하여 방제 의사결정에 도움이 될 것으로 생각된다. 나아가 본 연구에서 제시한 병원균의 감염위험 임계점을 사용하여 실제 포장에서 검증 실험을 수행함으로써 현장적용 가능성을 확인할 수 있을 것으로 생각된다.

## V. 참고문헌

- Carstensen, B., Plummer, M., Laara, E. and Hills, M. 2008. Epi: A package for statistical analysis in epidemiology. *R package version*. 2.32.
- Fabre, F., Dedryver, C. A., Leterrier, J. L. and Plantegenest, M. 2003. Aphid abundance on cereals in autumn predicts yield losses caused by *Barley yellow dwarf virus*. *Phytopathology*. 93: 1217-1222.
- Gillespie, T. J. and Sutton, J. C. 1979. A predictive scheme for timing fungicide applications to control *Alternaria* leaf blight in carrots. *Canadian Journal of Plant Pathology*. 1: 95-99.
- Harrell Jr, F. E., Harrell Jr, M. F. E. and Hmisc, D. 2019. Package ‘rms’. *Vanderbilt University*. 229.
- Hughes, G. 1999. Sampling for decision making in crop loss assessment and pest management: introduction. *Phytopathology*. 89: 1080-1083.
- Kim, K. W., Kim, K. R. and Park, E. W. 2016. Effects of interrupted wetness periods on conidial germination, germ tube elongation and infection periods of *Botryosphaeria dothidea* causing apple white rot. *The plant pathology journal*. 32: 1.
- Lee, J. H., Lee, J. M., Jung, H., Koo, B. S. and Jun, H D. 2017. Logistic Regression Model Construction and Evaluation for Maintenance of Water Level Stations. *Journal of the Korean Society of Hazard Mitigation*. 17: 389-396.
- Li, B. H., Xu, X. M., Li, J. T. and Li, B. D. 2005. Effects of temperature and continuous

- and interrupted wetness on the infection of pear leaves by conidia of *Venturia nashicola*. *Plant pathology*. 54: 357-363.
- Li, B. H., Yang, J. R., Dong, X. L., Li, B. D. and Xu, X. M. 2007. A dynamic model forecasting infection of pear leaves by conidia of *Venturia nashicola* and its evaluation in unsprayed orchards. *European journal of plant pathology*. 118: 227-238.
- Lindblad, M. and Waern, P. 2002. Correlation of wheat dwarf incidence to winter wheat cultivation practices. *Agriculture, ecosystems & environment*. 92: 115-122.
- MacHardy, W. E. and Gadoury, D. M. 1989. A revision of Mills' s criteria for predicting apple scab infection periods. *Phytopathology*. 79: 304-310.
- Madden, L. V. 2006. Botanical epidemiology: some key advances and its continuing role in disease management. *European Journal of Plant Pathology*. 115: 3-23.
- Mason, S. J. and Graham, N. E. 2002. Areas beneath the relative operating characteristics (ROC) and relative operating levels (ROL) curves: Statistical significance and interpretation. *Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography*. 128: 2145-2166.
- Mila, A. L., Driever, G. F., Morgan, D. P. and Michailides, T. J. 2005. Effects of latent infection, temperature, precipitation, and irrigation on panicle and shoot blight of pistachio in California. *Phytopathology*. 95: 926-932.
- Muller, M. P., Tomlinson, G., Marrie, T. J., Tang, P., McGeer, A., Low, D. E. and Gold, W. L. 2005. Can routine laboratory tests discriminate between severe acute respiratory syndrome and other causes of community-acquired pneumonia?. *Clinical infectious*

- diseases*. 40: 1079-1086.
- Nam, K. W., Han, M. K. and Yoon, D. H. 2014. Control effect of sodium dichloroisocyanurate for pear scab (*Venturia nashicola*) on Niitaka pear during flowering Period. *Korean Journal of Organic Agriculture*. 22: 347-357.
- Neufeld, K. N., Keinath, A. P. and Ojiambo, P. S. 2018. Evaluation of a Model for Predicting the Infection Risk of Squash and Cantaloupe by *Pseudoperonospora cubensis*. *Plant disease*. 102: 855-862.
- Pethybridge, S. J., Gent, D. H., Esker, P. D., Turechek, W. W., Hay, F. S. and Nutter Jr, F. W. 2009. Site-specific risk factors for ray blight in Tasmanian pyrethrum fields. *Plant disease*. 93: 229-237.
- R Core Team. 2017. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>
- Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J. C. and Müller, M. 2011. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC bioinformatics*. 12: 77.
- Sing, T., Sander, O., Beerenwinkel, N. and Lengauer, T. 2005. ROCR: visualizing classifier performance in R. *Bioinformatics*. 21: 3940-3941.
- Song, J.-H. and Seo, H.-J. 2018. Antifungal Activity of Agro-Materials against Pear Scab (*Venturia nashicola*) and Pear Rust (*Gymnosporangium asiaticum*) Fungi. *Research in Plant Disease*. 24: 33-40.
- Song, J. H., Lim, K. H., Cho, Y. S. and Lee, H. C. 2013. Current status of pest management



- and biodiversity in organic pear orchards in Korea. *Korean Journal of Organic Agriculture*. 21: 617-627.
- Suzuki, H. 1975. Meteorological factors in the epidemiology of rice blast. *Annual Review of Phytopathology*. 13: 239-256.
- Turechek, W. W. and Wilcox, W. F. 2005. Evaluating predictors of apple scab with receiver operating characteristic curve analysis. *Phytopathology*. 95: 679-691.
- Umemoto, S. 1991. Relationship between leaf wetness period, temperature and infection of *Venturia nashicola* to Japanese pear leaves. *Japanese Journal of Phytopathology*. 57: 212-218.
- Youden, W. J. 1950. Index for rating diagnostic tests. *Cancer*. 3: 32-35.
- Yuen, J., Twengström, E. and Sigvald, R. 1996. Calibration and verification of risk algorithms using logistic regression. *European Journal of Plant Pathology*. 102: 847-854.
- 농촌진흥청. 2016. 농작물 병해충 발생 예찰 요강
- 통계청. 2018. 농산물생산조사

## **ABSTRACT**

# **Development and evaluation of a pear scab forecast model using historical data**

**SOO YEON SEO**

Major in Plant Microbiology

Department of Agricultural Biotechnology

The Graduate School of Seoul National University

Plant disease occurrence varies depending upon weather conditions, and weather-based disease forecasts are often useful in decision-making for fungicide sprays. The decision-making process involves determining a threshold level of disease risk above which disease control action needs to be taken. In order to use disease forecast models practically for plant disease management, it is important to know the probability that disease forecast based on the risk threshold would be correct. In this study, we developed and evaluated a model that can be used to predict possible occurrence of pear scab epidemic prior to June 1 in a season. Pear scab disease incidence and weather data from 21 and 18 locations in 2014 and 2015, respectively, in Korea were used to develop the model, which consisted of a primary infection period after overwintering, an infection risk factor and an infection risk threshold. This model was evaluated using disease incidence

and weather data from 76, 69 and 57 locations in 2016 to 2018 in Korea. For developing the model, the Mills model which was originally developed to determine infection periods of apple scab was used to calculate daily infection risk ( $IR_d$ ). With the epidemiological characteristics of pear scab disease and historical disease data, we assumed that the primary infection period after overwintering were April 1<sup>st</sup> to May 16<sup>th</sup>, and pear scab outbreak prior to June 1<sup>st</sup> would cause significant damage to pear production of an orchard. The number of days with  $IR_d > 0$  during the period from April 1<sup>st</sup> to May 16<sup>th</sup> was applied to determine the probability of pear scab development with regard to the levels of infection risk threshold. The optimum infection risk threshold was determined to be 18 days by the receiver operating characteristic (ROC) curve analysis. Using the optimum infection risk threshold, accuracy of the model prediction on pear scab development was evaluated by the two-way contingency table analysis. The results indicated that the true positive warning rate (TPWR) and false negative warning rate (FPWR) of the disease forecast was 94 % and 17 % when the infection risk threshold was 18 days. In the validation study, the model showed that TPWR and FPWR were 70 % and 73 %, respectively.

**keywords** : Pear scab disease, Disease management, Forecast model, Decision-making, Binary logistic regression, ROC curve

**Student number** : 2017-27751